

Phylogenetic Distribution

The **Phylogenetic Distribution** option under **Compare Genomes** include 3 functions:

- Metagenomes vs. Genomes: shows reference isolate genome hits for selected metagenomes
- Genome vs. Metagenomes: shows the distribution of a reference genome in selected metagenome datasets
- Radial Tree: shows phylogenetic distribution of up to 5 genomes in a radial tree display

Metagenomes vs. Genomes

Metagenomes vs. Genomes comparison allows users to see the isolate reference genome composition of selected metagenomes. For example, a user can select 7 acid mine drainage metagenome datasets to check their composition (Figure 1(i)).

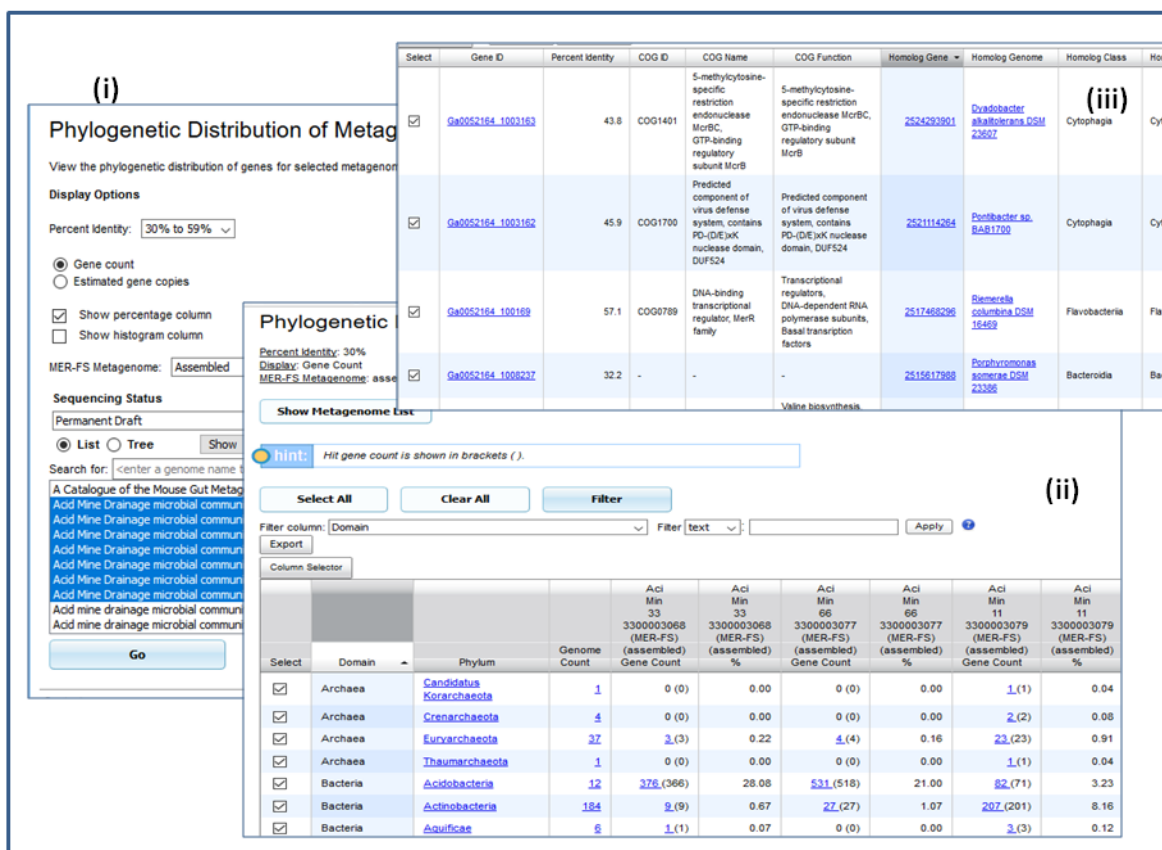


Figure 1. Metagenomes vs. Genomes.

The result (Figure 1(ii)) shows a row for each phylum. The table has the following fields:

- Domain: the domain of reference genomes
- Phylum: the phylum of reference genomes
- Genome Count: the number of reference genomes in the phylum showing up in the selected metagenome datasets

For each selected metagenome dataset, there will be 2 additional columns:

- Gene Count: with format $m(n)$, where m is the number of hits, and n is the count of distinct genes being hit; for example, 15 (9) means that there are 15 hits from this metagenome dataset to 9 distinct genes in the phylum
- Percentage: $= m / p$, where m is the number of hits in the previous column, and p is the total number of protein coding genes in the dataset

Users can click on the gene count to see the list of hit genes in a dataset (see Figure 1(iii)).

It is also possible to click on a Phylum in Figure 1(ii) to drill down to Class (Figure 2(i)), Order (Figure 2(ii)), all the way to Species level (Figure 2(iii)). In any of the pages, one can click on the **Show Metagenome List** button (highlighted in Figure 2(i)) to view the current metagenome dataset selection.

(iii)		Aci Min 33 3300003068 (MER-FS) (assembled)	Aci Min 66 3300003077 (MER-FS) (assembled)	Aci Min 11 3300003079 (MER-FS) (assembled)	Aci Min 44 3300003081 (MER-FS) (assembled)	Aci Min 77 3300003082 (MER-FS) (assembled)	Aci Min 55 3300003085 (MER-FS) (assembled)	Aci Min 22 3300003089 (MER-FS) (assembled)
Species	Genome Count	Gene Count	Gene Count	Gene Count	Gene Count	Gene Count	Gene Count	Gene Count
Methanocaldococcus jannaschii	1	0	0	2 (2)	1 (1)	0	1 (1)	0

(ii)		Aci Min 33 3300003068 (MER-FS) (assembled)	Aci Min 66 3300003077 (MER-FS) (assembled)	Aci Min 11 3300003079 (MER-FS) (assembled)	Aci Min 44 3300003081 (MER-FS) (assembled)	Aci Min 77 3300003082 (MER-FS) (assembled)	Aci Min 55 3300003085 (MER-FS) (assembled)	Aci Min 22 3300003089 (MER-FS) (assembled)
Order	Genome Count	Gene Count	Gene Count	Gene Count	Gene Count	Gene Count	Gene Count	Gene Count
Methanocellales	2	1 (1)	0 (0)	2 (3)	2 (3)	0 (0)	0 (0)	0 (0)
Methanomicrobiales	6	0 (0)	1 (1)	2 (3)	4 (4)	2 (2)	1 (1)	1 (1)
Methanosarcinales	7	0 (0)	1 (1)	5 (5)	2 (3)	2 (2)	0 (0)	1 (1)

Show Metagenome List

Filter column: Class

Export

Column Selector

(i)		Aci Min 33 3300003068 (MER-FS) (assembled)	Aci Min 33 3300003068 (MER-FS) (assembled)	Aci Min 66 3300003077 (MER-FS) (assembled)	Aci Min 66 3300003077 (MER-FS) (assembled)	Aci Min 11 3300003079 (MER-FS) (assembled)	Aci Min 11 3300003079 (MER-FS) (assembled)
Class	Genome Count	Gene Count	%	Gene Count	%	Gene Count	%
Archaeoglobi	2	0 (0)	0.00	0 (0)	0.00	2 (2)	0.08
Halobacteria	4	0 (0)	0.00	0 (0)	0.00	4 (4)	0.16
Methanobacteria	1	0 (0)	0.00	0 (0)	0.00	1 (1)	0.04
Methanococci	5	2 (2)	0.15	2 (2)	0.08	1 (1)	0.04
Methanomicrobia	16	1 (1)	0.07	2 (2)	0.08	11 (11)	0.43
Thermococci	6	0 (0)	0.00	0 (0)	0.00	2 (3)	0.12
Thermoplasmata	1	0 (0)	0.00	0 (0)	0.00	0 (0)	0.00
Unclassified	1	0 (0)	0.00	0 (0)	0.00	1 (1)	0.04

Export

Figure 2. Genome Hit Drill Down.

Genome vs. Metagenomes analysis allows user to check whether an isolate genome is in the composition of a set of selected (up to 5000) metagenome datasets.

For example, a user can select *Methanoregula boonei* 6A8 (IMG OID: 640753014) to be the reference isolate genome, and select 8 acid mine drainage metagenome datasets to run the comparison using **Group hits by phylum** output type (see Figure 3(i)).

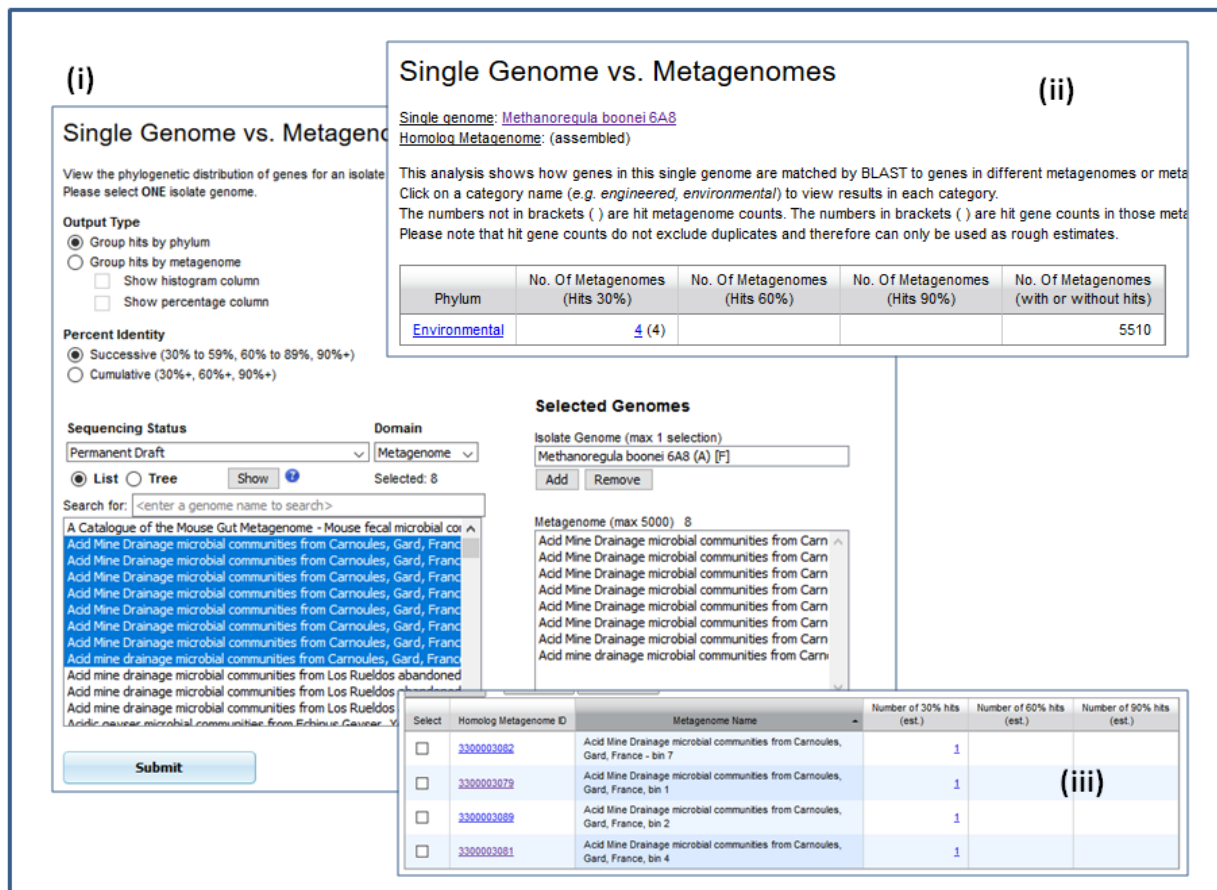


Figure 3. Genome vs. Metagenomes.

The result in Figure 3(ii) shows four metagenome datasets have hits. Detailed information of the 4 datasets can be examined by clicking on the number 4 in Figure 3(ii). The four metagenome datasets are listed in Figure 3(iii).

Like the Metagenomes vs. Genomes option, users can also drill down on Phylum in Figure 3(ii) to view Class (Figure 4(ii)), Order (Figure 4(ii)), all the way down to Species level (Figure 4(iii)).

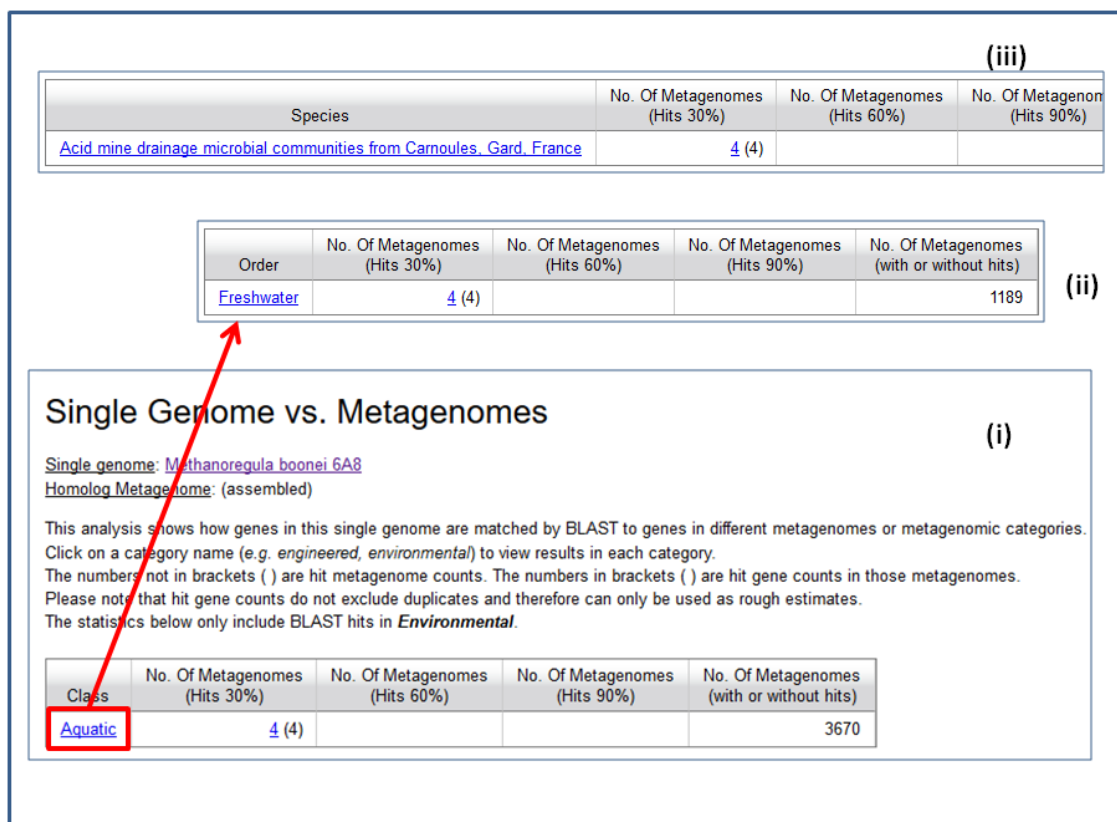


Figure 4. Genome vs. Metagenomes Drill Down.

If the user selects **Group hits by metagenome** output type option (Figure 3(i)), then the result will be displayed with each selected metagenome having hits in a separate row as shown in Figure 5(i). Click on metagenome name to see the detailed information of the metagenome (Figure 5(ii)), and click on hit count to see the list of hits (Figure 5(iii)).

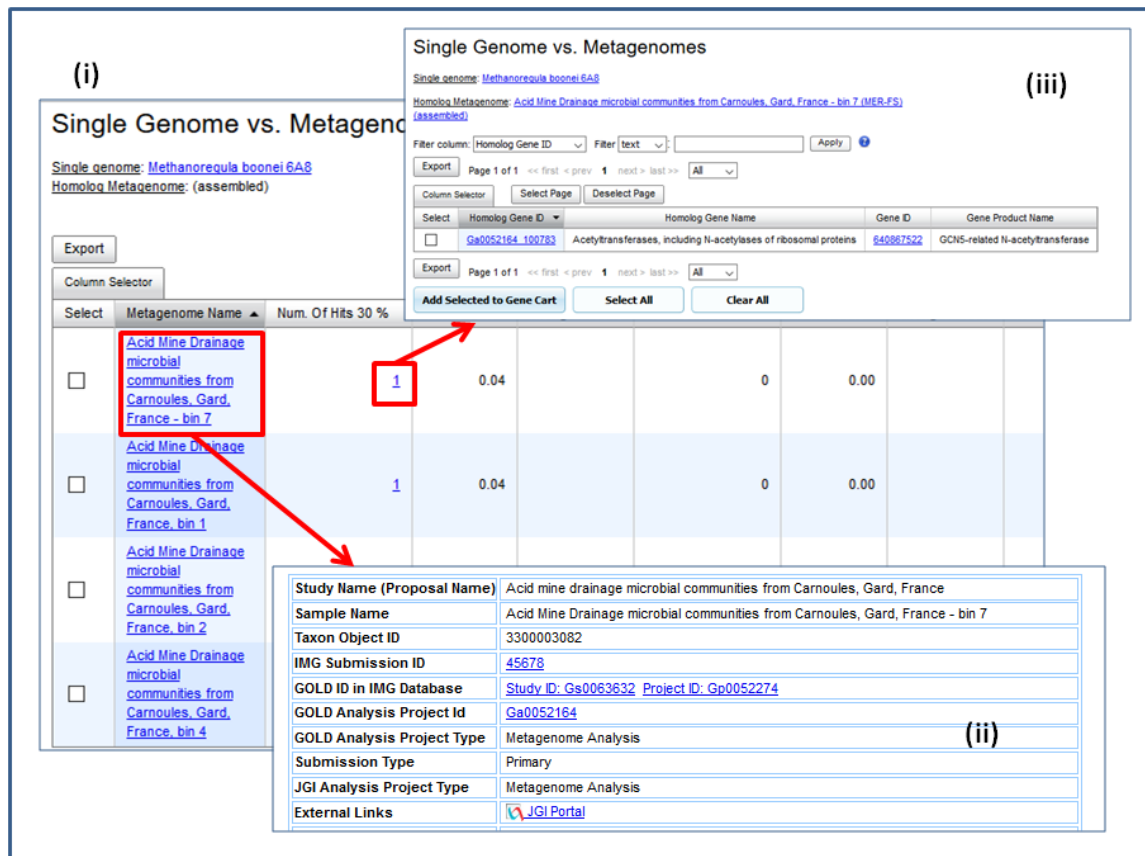


Figure 5. Genome vs. Metagenomes -- Group hits by metagenome

Radial Tree

Users have the ability to view up to five genomes phylogenetically in a circular tree. This tool is found under menu Compare Genomes - Radial Tree, see Figure 6.

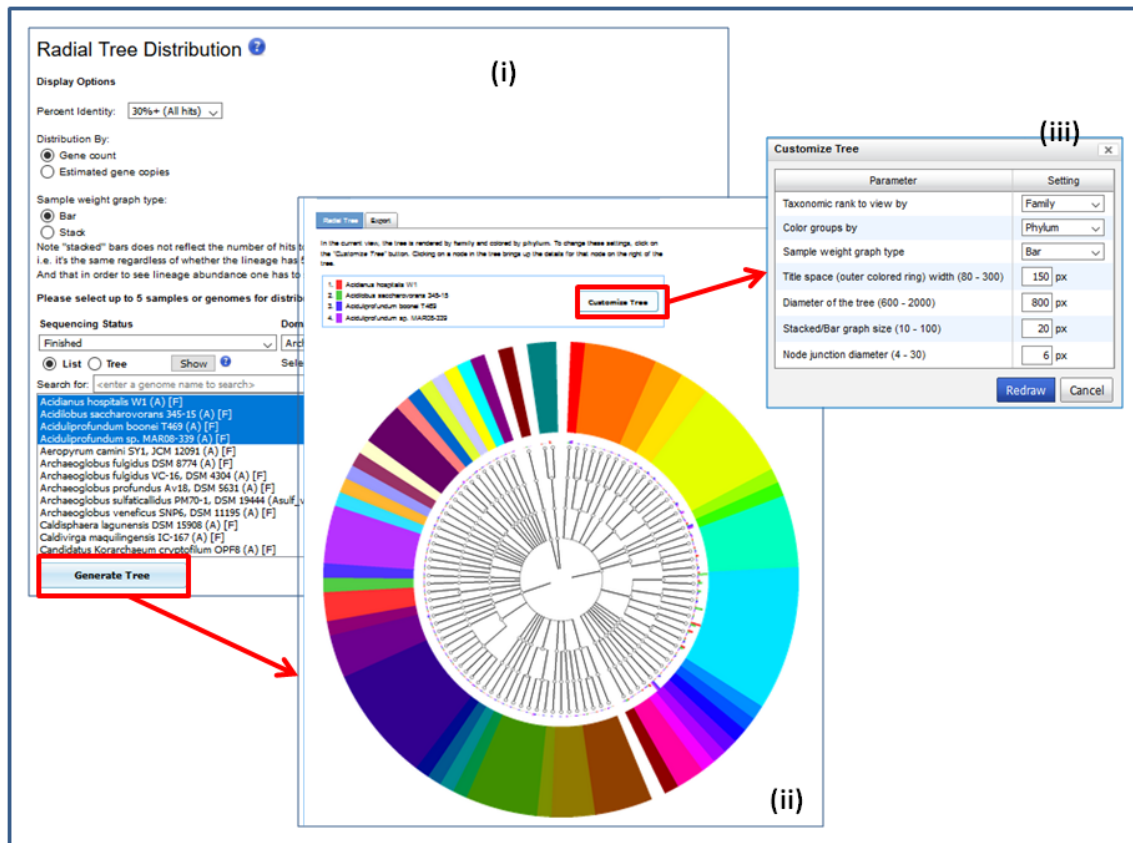


Figure 6. Radial Tree Distribution.

For example, a user selects 4 archaeal genomes and click the **Generate Tree** button (Figure 6(i)) to view the radial tree distribution of the 4 genomes (Figure 6(ii)). The tree display can be customized by clicking **Customize Tree** button in Figure 6(ii). Customization options include:

- Taxonomic rank / tree level by: Domain, Phylum/Class, Order and Family
- Color group by: Domain, Phylum/Class, Order and Family
- Genome sample weight graph: Stacked, Bar
- Title space
- Diameter of the tree
- Stacked/Bar graph size
- Node junction diameter

Credit: Radial Phylogenetic Tree tool supplied with the help of:

The Metagenomics RAST server - A public resource for the automatic phylogenetic and functional analysis of metagenomes F. Meyer, D. Paarmann, M. D'Souza, R. Olson, E. M. Glass, M. Kubal, T. Paczian, A. Rodriguez, R. Stevens, A. Wilke, J. Wilkening, R. A. Edwards
BMC Bioinformatics 2008, 9:386